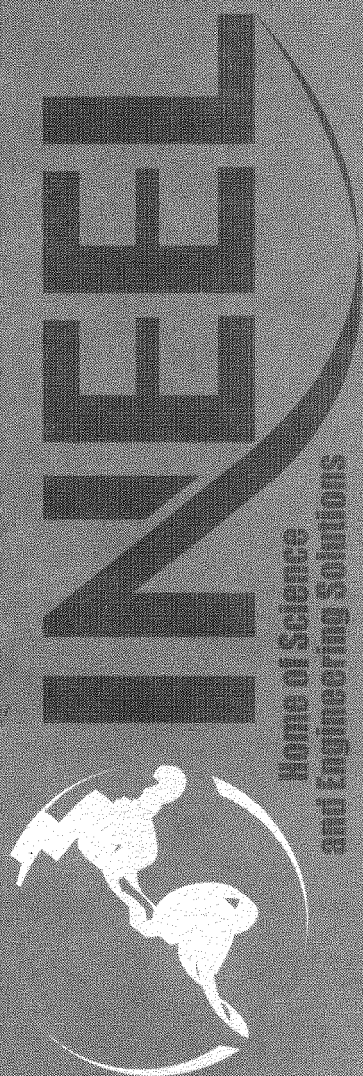


Geostatistic Modeling of Subsurface Characteristics in the Radioactive Waste Management Complex Region, Operable Unit 7-13/14

*Molly K. Leecaster
April 2002*



*Idaho National Engineering and Environmental Laboratory
Bechtel BWXT Idaho, LLC*

Geostatistic Modeling of Subsurface Characteristics in the Radioactive Waste Management Complex Region, Operable Unit 7-13/14

Molly K. Leecaster

April 2002

**Idaho National Engineering and Environmental Laboratory
Environmental Restoration Program
Idaho Falls, Idaho 83415**

**Prepared for the
U.S. Department of Energy
Assistant Secretary for Environmental Management
Under DOE Idaho Operations Office
Contract DE-AC07-99ID13727**

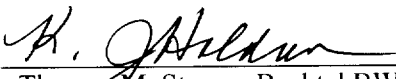
Geostatistic Modeling of Subsurface Characteristics in the Radioactive Waste Management Complex Region, Operable Unit 7-13/14

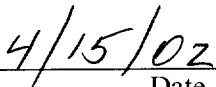
INEEL/EXT-02-00029

Revision 0

April 2002

Approved by


for Thomas M. Stoops, Bechtel BWXT Idaho, LLC
Waste Area Group 7 Project Engineer


Date

ABSTRACT

This report provides documentation of an effort to predict values based on available sample site information for variables in lithology, hydraulic characteristics, and aquifer permeability data sets for the vadose zone and the Snake River Plain Aquifer in the Radioactive Waste Management Complex region of the Idaho National Engineering and Environmental Laboratory. This work was performed in support of the Operable Unit 7-13/14 remedial investigation and feasibility study.

CONTENTS

ABSTRACT.....	v
1. INTRODUCTION.....	1
1.1 Purpose	1
1.2 Scope of Work	1
1.3 Spatial Models.....	4
1.4 Background.....	4
2. DATA.....	6
2.1 Lithology Data Set	6
2.1.1 Irregularities in the Lithology Data.....	6
2.1.2 Predicting Elevation of Subsurface Layers and Thickness	7
2.2 Hydraulic Characteristics Data Set.....	8
2.3 Aquifer Permeability Data Set	8
3. MODELS.....	9
3.1 Inverse Distance Weighting	9
3.2 Simple Kriging	9
3.3 Universal Kriging	11
3.4 Variograms.....	11
4. IMPLEMENTATION	13
4.1 Grids	13
4.2 Lognormal Data	13
4.3 Effects of Sample Location.....	13
4.4 Model Fit Assessment	14
5. RESULTS.....	15
5.1 Lithology Data Set	15
5.2 Hydraulic Characteristics Data Set.....	20
5.3 Aquifer Transmissivity Data Set.....	21

6.	DISCUSSION	22
6.1	Precision of the Predictions and Fit of the Model	22
6.2	Cross Validation	23
6.3	Sparseness of Data	23
6.4	Improvements to Predictions	24
7.	REFERENCES.....	25
Appendix A—Raw Data in Lithology, Hydraulic Characteristics, and Aquifer Permeability of the Vadose Zone.....		A-1
Appendix B—Variograms and Maps		B-1

FIGURES

1.	Map of the Radioactive Waste Management Complex at the Idaho National Engineering and Environmental Laboratory	2
2.	Map of the Subsurface Disposal Area at the Radioactive Waste Management Complex	3
3.	Prediction domains over the Subsurface Disposal Area and surrounding territory	4
4.	Examples of a basalt gap between the subsurface soil layer and the A-B interbed soil layer and an interbed gap between the B and C basalt layers	7
5.	Directional variograms for the C basalt layer raw data (top) and residuals from regression on easting and easting ² (bottom)	16
6.	Difference between prediction of bottom elevation of A-B interbed soil from Methods 1 and 2 for dealing with gaps	17

TABLES

1.	Correlation among variables from the lithology data set.....	5
2.	Summary of stationarity investigation showing significant ($p < 0.1$) regression ($\sqrt{}$) of variables on directions east and north, their squares, and the interaction	10
3.	Interpolation method and variogram parameters (where appropriate) used to model basalt layer elevations, interbed soil thickness, soil permeability and porosity, and aquifer permeability	17
4.	Comparison of data evaluation techniques for characterizing lithologic layers (units in feet)	18
5.	Cross-validation statistics for lithology data	19
6.	Cross validation statistics for hydraulic characteristics	21
7.	Cross validation statistics for aquifer permeability based on log-transformed data	21
8.	Modeling standard error for Grid Refinement 2	22

Geostatistic Modeling of Subsurface Characteristics in the Radioactive Waste Management Complex Region, Operable Unit 7-13/14

1. INTRODUCTION

The *Federal Facilities Agreement and Consent Order for the Idaho National Engineering Laboratory* (DOE-ID 1991) subdivides the Idaho National Engineering and Environmental Laboratory (INEEL) into 10 waste area groups (WAGs) for remedial management. Waste area group is also the designation recognized under the Comprehensive Environmental Response, Compensation and Liability Act (CERCLA) (42 USC § 9601 et seq.). The Radioactive Waste Management Complex (RWMC) at the INEEL (see Figure 1) was designated as WAG 7. This geostatistic modeling study was performed for the Subsurface Disposal Area (SDA), shown in Figure 2, which is part of the RWMC.

Hydrologic models are important tools for predicting fate and transport of groundwater and the constituents carried therein. Understanding flow mechanisms is especially important for areas like the SDA where hazardous constituents may be in the groundwater. To apply deterministic hydrologic models, many characteristics need to be known at a small scale over the entire area of interest. Characteristics for the area of concern include the following:

- Depth of subsurface soil and basalt layers (lithology)
- Porosity and permeability of the interbed (hydraulic characteristics)
- Permeability of the Snake River Plain Aquifer.

These variables are measured at only a sample of sites. Predictions must be made because hydrologic modelers seek to determine values over the whole area of interest.

1.1 Purpose

This report provides documentation of an effort to predict values for variables in lithology, hydraulic characteristics, and aquifer permeability data sets (see Appendix A) for the area in and around the SDA, based on available sample site information. This work supports the CERCLA-driven Operable Unit 7-13/14 remedial investigation and feasibility study.^a

1.2 Scope of Work

The geologic and hydrologic characteristics data are used to determine appropriate spatial models and calculate predictions. This process involves a close exploration of the data, an investigation of anisotropy (directionally dependent spatial correlation), consideration of extreme values, examination of relationship to location variables, calculation of predictions, determination of the fit of the model, and

a. Holdren, K. Jean, Bruce H. Becker, Nancy L. Hampton, L. Don Koeppen, Swen O. Magnuson, T. J. Meyer, Gail L. Olson, and A. Jeffrey Sondrup, 2002, "Waste Area Group 7, Operable Unit 7-13/14, Comprehensive Remedial Investigation/Feasibility Study (Draft)," DOE/ID-10834, Rev. B, U.S. Department of Energy Idaho Operations Office, Idaho Falls, Idaho, March 2002.

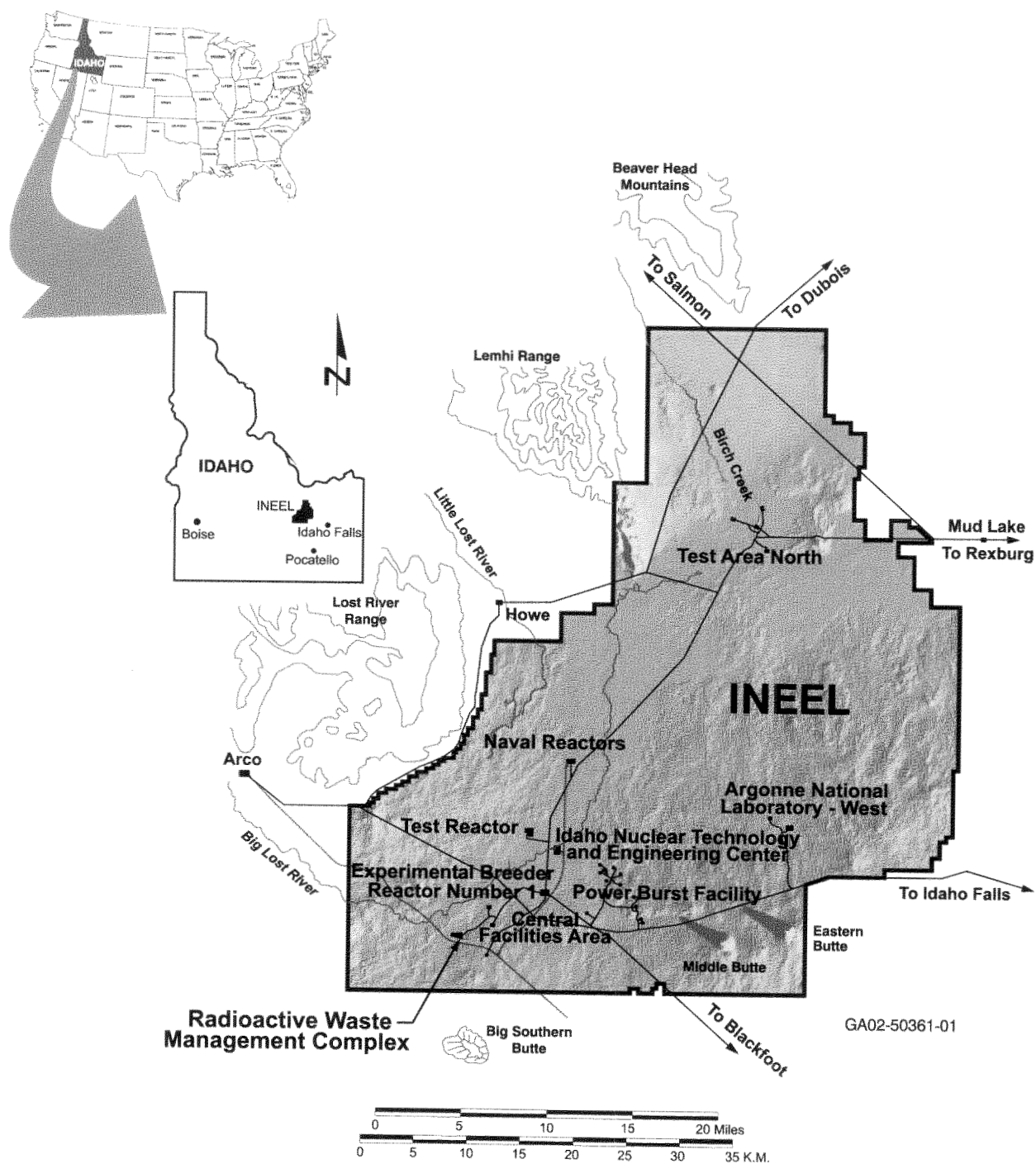


Figure 1. Map of the Radioactive Waste Management Complex at the Idaho National Engineering and Environmental Laboratory.

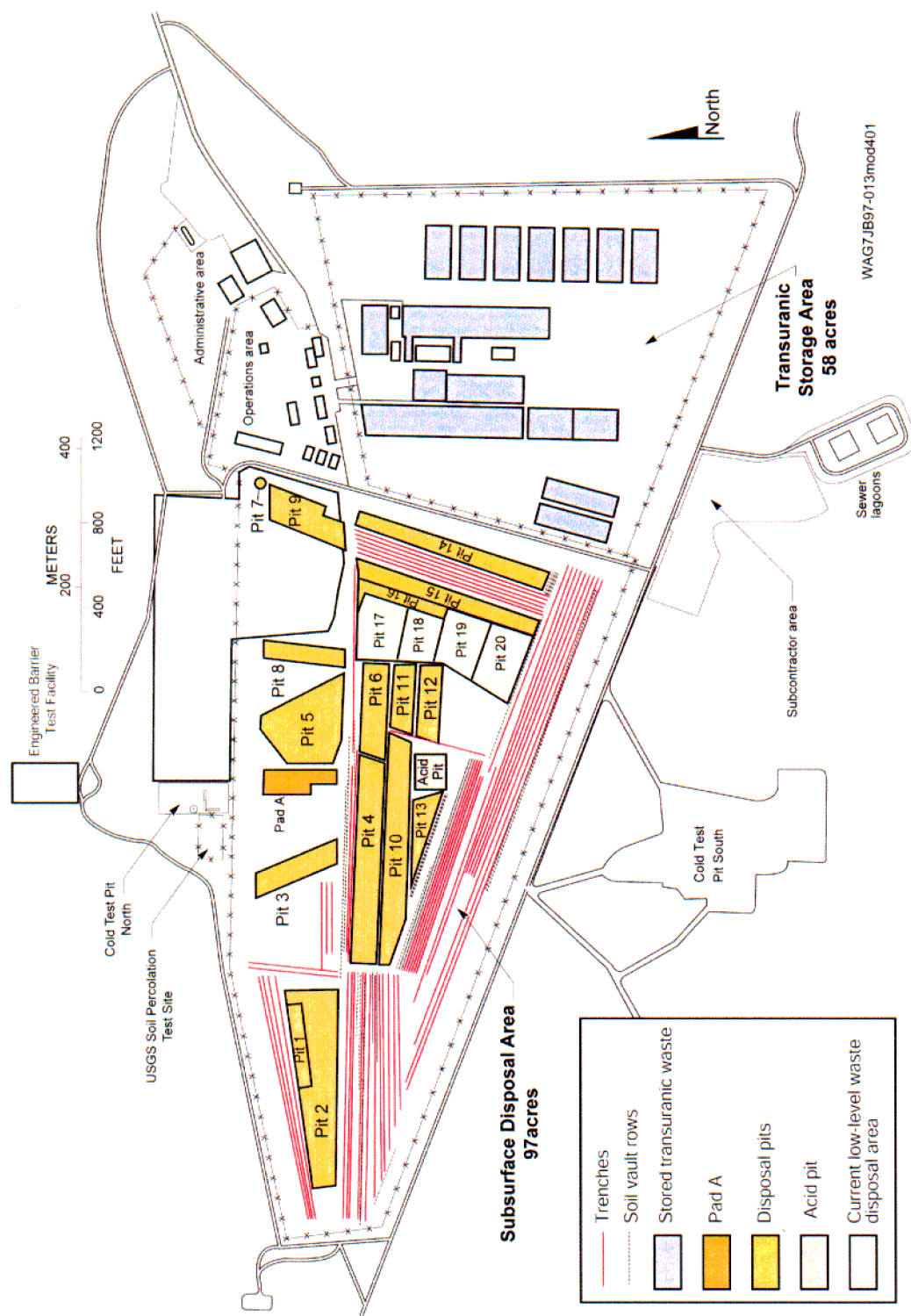


Figure 2. Map of the Subsurface Disposal Area at the Radioactive Waste Management Complex.

examination of the model variability. These steps are performed for each of the three data sets (i.e., lithology, hydraulic characteristics, and aquifer permeability) along with specific analyses pertaining to each data set.

1.3 Spatial Models

Three spatial models are used to predict values over the area: (1) inverse distance weighting, (2) simple kriging, and (3) universal kriging. The predictions are made for four domains or grid refinements. These grid refinements are rectangular and of decreasing size and increasing intensity (see Figure 3). This follows roughly the pattern of sample intensity as well as data needs for hydrologic modeling. Benefits of the methods are described and resulting predictions are presented in the following sections.

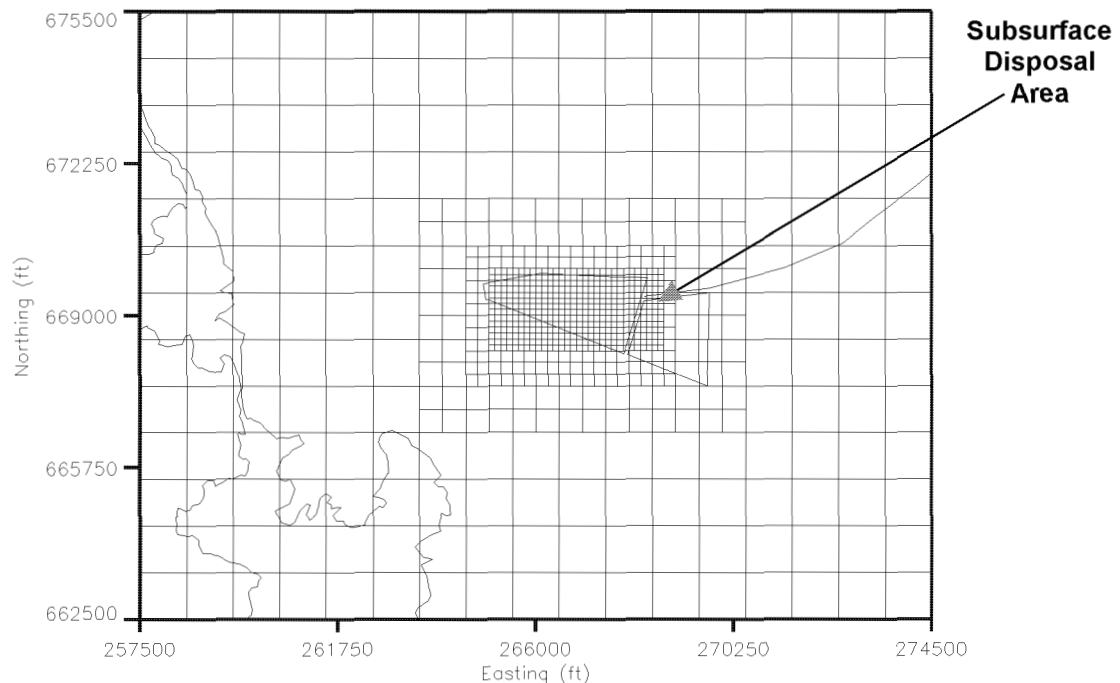


Figure 3. Prediction domains over the Subsurface Disposal Area and surrounding territory.

1.4 Background

The geologic characteristics of the area are described in *Development, Calibration, and Predictive Results of a Simulator for Subsurface Pathway Fate and Transport of Contaminants in the SDA* (Magnuson and Sondrup 1998). The data in Table 1 are based on *Stratigraphic Data for Wells At and Near the Idaho National Engineering Laboratory, Idaho* (Anderson et al. 1996) with interpretations for basalt and sediment interfaces derived from data on wells drilled in 1999. The hydraulic characteristics were compiled from four sources (Barraclough et al. 1976; McElroy and Hubbell 1990; correspondence from Southwest Research Institute;^b and USGS 2000, p. 30). The aquifer permeability data were combined from *Aquifer Testing of Wells M1S, M3S, M4D, M6S, M7S, and M10S at the Radioactive Waste Management Complex* (Wylie and Hubbell 1994) and *Pumping Test of Pit 9 Production Wells* (Wylie 1996).

b. Mike Dammann, Southwest Research Institute, Letter to Corey Frandsen, Idaho National Engineering and Environmental Laboratory, November 2, 2000, "Idaho National Engineering and Environmental Laboratory Scope of Work," ER-SOW-371.

Table 1. Correlation among variables from the lithology data set.^a

Layer	C-D Interbed Soil	C Basalt	B-C Interbed Soil	B Basalt	A-B Interbed Soil	A Basalt
Surface elevation	-0.06	-0.17	0.30	0.03	-0.26	0.76
Surficial soil	-0.16	0.00	-0.17	-0.12	-0.20	-0.26
A basalt	-0.04	-0.36	0.38	-0.14	-0.46	NA
A-B interbed soil	0.02	0.38	-0.13	-0.03	NA	NA
B basalt	0.40	0.15	-0.22	NA	NA	NA
B-C interbed soil	0.02	-0.42	NA	NA	NA	NA
C basalt	0.03	NA	NA	NA	NA	NA

a. Values in **bold** font are statistically ($\alpha = 0.05$) different from zero.

NA = not applicable

Note: Only interbed soil-layer information is useful for estimating truncated values. Other values are included for completeness.

2. DATA

Three separate data sets were explored and used to calibrate spatial models: (1) lithology, (2) hydraulic characteristics, and (3) aquifer permeability. Sample site locations were measured in state plane coordinates in the North American Datum of 1927 (i.e., NAD 27), which is historically used on the INEEL. The following sections contain descriptions of the data sets, an explanation of the models, and some general steps that were involved in the modeling process.

2.1 Lithology Data Set

The lithology data set (see Appendix A, Table A-1) consisted of 118 sample sites where bore holes were drilled and elevation measurements taken at the bottom of the basalt layers and thickness of the interbed soil layers. Eighteen layers were measured, but only eight were used in the analyses documented in this report. These layers are listed below:

- Surface elevation
- Surficial soil thickness
- Elevation of the bottom of the A basalt layer
- Thickness of the A-B interbed soil layer
- Elevation of the bottom of the B basalt layer
- Thickness of the B-C interbed soil layer
- Elevation of the C basalt layer
- Thickness of the C-D interbed soil layer.

Because the amount of unrecorded data increased as the depth increased, the distance between sites in the lower layers (i.e., D basalt layer and D-E interbed soil layer) exceeded the spatial continuity of the data. The D basalt layer and D-E interbed soil layer, as well as layers below, had too few sampled sites to consider modeling.

2.1.1 Irregularities in the Lithology Data

The lithology data had three irregularities: (1) truncated values, (2) gaps in soil or basalt layers, and (3) variables measured on different scales. These were dealt with in a variety of ways.

2.1.1.1 Truncated Values. First, truncated values resulted when boreholes were drilled into but not through a soil layer. These were dealt with by either using the reported values or an estimate. The value could be estimated from regression on upper layer elevation and thickness values because some of these variables were moderately correlated (see Table 1).

2.1.1.2 Gaps in Soil or Basalt Layers. The second irregularity in the data was gaps in the layers produced from merging soil layers (i.e., basalt gap) or merging basalt layers (i.e., interbed gap) (see Figure 4). These gap situations were dealt with using one of three methods. In Method 1 (the default method used by Magnuson and Sondrup [1998]), the two types of gaps (i.e., interbed and basalt) were

Environment

Soil

Basalt



Figure 4. Examples of a basalt gap between the subsurface soil layer and the A-B interbed soil layer and an interbed gap between the B and C basalt layers.

treated the same. The lower soil layer of a basalt gap was assigned a thickness of zero and the upper soil layer was assigned the whole thickness. When two basalt layers merged and an interbed gap existed, the data already reflected the correct basalt layer bottom elevations. Therefore, no alteration resulted and the soil layer had a thickness of zero. Methods 2 and 3 treated the two gap types differently. Where basalt layers merged (i.e., an interbed gap), the lower soil layer was assigned the correct thickness of zero for both approaches. For Method 2 (i.e., in basalt gap areas), the two soil layers were assigned the same thickness (i.e., that of the merged depth). For Method 3, each soil layer was assigned one-half of the merged depth thickness.

2.1.1.3 Variables Measured on Different Scales. The third irregularity was that the data were alternately reported in thickness and elevation. These variables had very different value ranges and also made every second layer independent of the upper layers. Three data sets were constructed to determine the effect of this irregularity. The first was the reported data in alternating thickness and elevation. All values were converted into thickness in the second data set and all values were converted into elevations in the third data set.

2.1.2 Predicting Elevation of Subsurface Layers and Thickness

Regardless of what data set was used for analysis, in the final presentation all data were reported as elevation at the bottom of the layer. Therefore, the surficial soil elevation was also the elevation at the top of the A basalt layer. With two alternatives for truncated values, three alternatives for gaps, and three alternatives for data units, 18 data setups were considered and compared in predicting elevation of the subsurface layers.

The 18 data setups were compared by calculating the differences between predictions at all grid points. The mean difference between predictions from the data setups was calculated. Mean absolute differences also were computed to determine whether the differences were always one-way or whether sometimes one method gave lower and sometimes higher predictions than another method. These gave an overall comparison among setups. If differences were small, then which setup was used did not matter and the raw data were preferred for simplicity. In some instances, the differences were mapped to see

where large differences occurred and if they made sense. For instance, large differences were expected between the two methods for dealing with sites that include truncated values. Geologists explored setups that produced different predictions to determine which representation was most likely to reflect reality.

The lithology layers did not overlap in reality, but predicted layers did. In predicting values for the lithology layers, a separation constraint was imposed to ensure that the layers did not overlap, but did account for them possibly coinciding. The first step was to change negative thickness predictions to zero. After the predictions were changed to elevations, they were checked for overlap. If layers overlapped, the lower layer value was set equal to the upper layer.

2.2 Hydraulic Characteristics Data Set

The hydraulic characteristic data set (see Appendix A, Table A-2) consisted of interbed permeability or porosity for a total of 112 samples. These samples represented 32 site locations at 1 to 19 different depths. The data covered the B-C and C-D soil interbeds. Data for the B-C interbed were obtained from 17 sites and data for the C-D interbed from 24 sites (see Appendix A, Table A-3). Permeability was measured in centimeters per second and porosity was measured in percent. The porosity data from various depths at each site location were combined by taking the mean. The permeability data were transformed to millidarcies before taking the harmonic mean over depth for each site location. Harmonic averaging was used because it is most appropriate for flow perpendicular to layering (Freeze and Cherry 1979, pp. 33 to 34). Flow perpendicular to layering is thought to be most appropriate in the case of the sedimentary interbeds that are horizontally deposited and primarily have vertical water movement through them. The harmonic mean was then natural log transformed to attain symmetry.

2.3 Aquifer Permeability Data Set

The aquifer permeability data set (see Appendix A, Table A-4) consisted of 21 sample sites with transmissivity measured in cubed feet per day per foot. These data were natural log transformed to attain symmetry. Four values outside the base grid were influential in the predictions. These sites were removed because they were farther away from base grid sites than were other sites within the base grid.

3. MODELS

The three different modeling techniques of inverse distance weighting, simple kriging, and universal kriging were used. The final technique was chosen depending on its ability to accurately achieve the goal of producing maps that provided a realistic picture over the area of interest. Kriging takes into account the spatial relationship of a series of points, permitting interpolation within a three-dimensional matrix of points and estimation of additional data points. While the kriging methods had the added benefit of a derived measure of precision as well as some information about spatial continuity, they also had the added burden of modeling the variogram, which required adequate data at appropriate distances.

3.1 Inverse Distance Weighting

One modeling technique applied to data was inverse distance weighting. This technique made no assumptions but could provide an ad hoc measure of precision (Reich and Davis 2000). The prediction at a grid point (\hat{z}_i) was calculated by weighting each measured value (x_j) by the inverse of the distance between the prediction site and the sample site ($d_{i,j}$), as follows:

$$\hat{z}_i = \frac{\sum_{j=1}^n x_j / d_{i,j}}{\sum_{j=1}^n 1/d_{i,j}} \quad (1)$$

where

\hat{z}_i = predicted value

x_j = measured value

$d_{i,j}$ = distance between the prediction site and sample site.

The variance of the prediction (\hat{v}) was given by

$$\hat{v}(\hat{z}_i) = \left(\frac{n}{n-1} \right) \frac{\sum_{j=1}^n \left(x_j / d_{i,j} - \hat{z}_i / d_{i,j} \right)^2}{\left(\sum_{j=1}^n 1/d_{i,j} \right)^2} \quad (2)$$

Inverse distance weighting was the selected method for the hydraulic characteristic interbed permeability for the B-C and C-D soil interbeds. This method was chosen because of inadequate data for proper variogram modeling. This method was implemented using 3DField software (Galouchko 2002).

3.2 Simple Kriging

The second modeling technique applied to data was simple kriging, which made predictions based on a weighted mean of all sample values. The two assumptions made in kriging were (1) of a locally common expected value (i.e., stationarity) and (2) that spatial correlation was independent of direction

(i.e., isotropy). Stationarity was assumed, and in simple kriging the mean function was assumed known. Isotropy of all variables was verified by comparing the variograms (i.e., values of distance and variance between pairs of points at each distance [see Appendix B, Figure B-1]) of the raw data and of the residuals from a polynomial regression on related terms specified in Table 2. These variograms were compared for four directions corresponding to northeast-southwest, north-south, northwest-southeast, and east-west. The directional variogram was calculated using only pairs that were in the specified direction, plus or minus a tolerance angle. The data sets that displayed similar variograms for all directions were considered isotropic. The empirical variograms of the isotropic data then were modeled by specifying each of the following:

- **Model:** Gaussian or spherical
- **Nugget:** small scale variation
- **Sill:** asymptotic variance between independent pairs of sites
- **Range:** distance at which the sill is attained.

Table 2. Summary of stationarity investigation showing significant ($p < 0.1$) regression (\checkmark) of variables on directions east and north, their squares, and the interaction.

Data Set	Variable	East	North	East ²	North ²	East x North
Lithology	Surface elevation	\checkmark		\checkmark		
	Surficial soil thickness					
	A basalt elevation					
	A-B interbed soil thickness		\checkmark		\checkmark	
	B basalt elevation		\checkmark		\checkmark	
	B-C interbed soil thickness					
	C basalt elevation	\checkmark		\checkmark		
	C-D interbed soil thickness	\checkmark		\checkmark		
Interbed hydraulic	B-C interbed porosity	\checkmark	\checkmark			\checkmark
	C-D interbed porosity	\checkmark				
	B-C interbed permeability					
	C-D interbed permeability		\checkmark		\checkmark	
Aquifer permeability	Transmissivity	\checkmark	\checkmark			

Kriging predictions were calculated to minimize the error variance, which was accomplished by weighting station values in accordance with the variogram model parameters. Those values closer to the prediction point received larger weight. Because values from sites close to the prediction sites were more highly correlated with the prediction, the variance was smaller when sites were close together. Simple kriging was used for the following lithology variables:

- Surficial soil thickness

- A basalt elevation
- B-C interbed soil thickness.

Simple kriging was performed in S+SPATIALSTATS (MathSoft 2000).

3.3 Universal Kriging

The third modeling technique applied to data was universal kriging, which incorporated the assumption that the data have an unknown mean function. The procedure was the same as simple kriging except that the mean function was estimated as a function of easting and northing in the process of calculating predictions. Terms possibly included were easting, northing, easting², northing², and easting x northing. Only terms that were significantly related to the response were included. Extra terms that were not useful in prediction increased the kriging variance unduly. Estimates of percent of variability explained by the linear and polynomial trend were obtained while performing these regressions to determine which terms to include in the model. This assisted in the selection of terms and also determined the utility of universal kriging over simple kriging. A large part of the remaining variability (after the terms were included) was assumed to be explained by the spatial model. Universal kriging was applied to the remaining lithology variables as well as the interbed soil porosity and aquifer transmissivity. Universal kriging also was performed in S+SPATIALSTATS (i.e., Splus).

3.4 Variograms

Simple or universal kriging began by modeling the empirical variogram of the isotropic data. Various parameters were specified to ensure robustness in the calculation of the variogram. The parameters are listed below:

- Number of distances
- Size and tolerance of the distance
- Maximum distance between pairs
- Estimation type (classical or robust [Cressie 1993])
- Minimum number of pairs used to estimate at each distance.

After developing an appropriate empirical variogram, a model was fit. Manual model fitting of the variogram included the following steps:

- Trying more than one model
- Trying various values for sill, nugget, and range
- Analyzing variogram clouds
- Calculating cross-validation variograms
- Comparing kriging estimates from different variogram models.

Some of these data sets had few pairs of site locations within 305 m (1,000 ft), which made variogram modeling difficult. To determine the relationship between pairs of sites at close distances, a variogram cloud was considered. The variogram cloud portrayed all semivariances at all distances, not just an average (or median for robust calculation) for all pairs within a range of distances. This portrayal allowed a closer inspection of the relationship. Small-distance relationships could be discerned as well as the distribution of semivariances at all distances. The variogram cloud could be portrayed as a plot of all possible points. However, this plot became cluttered and difficult to interpret, especially in the presence of some extreme semivariances. The boxplots of values at many distances were more revealing because they displayed the median and quartiles of the semivariance distributions and downplayed the effect of the extreme values. In modeling the empirical variograms, the boxplots of variogram clouds also were considered (see Appendix B, Figure B-1).

To investigate single outliers, variograms were calculated with a hold-one-out method and compared. Each variogram was calculated after deleting one site and was compared to the variogram from the full data set. Variograms that differed from the rest indicated a problem with that site.

4. IMPLEMENTATION

Spatial modeling proceeded in steps. First was exploratory data analysis. Second was variogram calculation and modeling. The first step involved determining the best empirical variogram for raw data and residuals, looking at directional variograms and a variogram cloud, and modeling the resulting isotropic data. The next step was to calculate the model predictions for the all techniques. The predictions were made for each of four grids, which are described below. For data that were lognormal, the predictions were back-transformed to their original scale (also described below). Finally, the models were compared with respect to predictions and their standard deviations. The effect of sample location is discussed below as well as methods to compare predictions. Variability was examined using contour plots of model standard deviations for kriging and inverse distance weighting. These maps provided an indication of confidence in the model predictions (i.e., the larger the standard deviation, the less confidence in the model predictions). The model predictions are always more precise around sample sites and in areas of dense sampling. Model predictions are less reliable in areas of sparse sampling and have larger model standard deviations. Model standard deviations also are interpreted relative to the prediction values because larger values generally have larger variances.

4.1 Grids

The predictions were made on four grid refinements (see Figure 3). The four grids were devised to cover a large region with a coarse grid and successively smaller areas with finer grids. The results then would be combined so that those smaller-area grid (i.e., interior) predictions had precedence over larger-area grid predictions, thus creating a map of the whole area with different grid intensities in different areas.

4.2 Lognormal Data

Some of the data were lognormally distributed and were natural log transformed before variogram modeling and kriging. The distribution of the data did not impact inverse distance weighting; therefore, raw data were used there. To obtain estimates of the original data after kriging, the log-transformed data required a bias correction to the simple exponential back-transformation. This bias correction was plus one-half of the kriging variance. The back-transformation often resulted in predicted values that were at least an order of magnitude larger than the largest measured value. This effect was exacerbated by sparse data because the kriging variance was larger for data that were spread farther apart. Alternative approaches for lognormal data were attempted. The alternatives were to use (1) a different transformation (e.g., normal score), (2) indicator kriging, (3) simulation, or (4) another method such as inverse distance weighting. In practice, the normal score transformation produced nugget-only variograms. Indicator kriging required specific limits to compress the data into 0:1 data. Simulation was a viable alternative (as discussed in Section 6) and inverse distance weighting, which was straightforward and applicable, was used here.

4.3 Effects of Sample Location

The sample sites were not on a regular grid, nor were they evenly dispersed over the area for any of the data sets. The area in the SDA had many sample sites, with fewer and fewer as the distance away from the SDA increased. This distribution had an impact on all types of predictions. For kriging models, the distance between pairs of sites influenced the empirical variogram and hence the variogram model, which determined the kriging weights. The small distance range of the variogram model greatly affected the kriging predictions and that lack of data led to uncertain predictions. This uncertainty was only partially reflected in the kriging variance. The kriging variance was larger when data points were farther apart, but

this still understated the uncertainty associated with unsupported small-scale variogram modeling. In some instances, depending on how sparse the data were, kriging models were not appropriate. If the variogram could not be modeled sufficiently, then the kriging variance (a major advantage of kriging) was suspect, and simpler approaches such as inverse distance weighting were preferred. The inverse distance weighting predictions also were somewhat impacted because the closest sites, even if they were far away, still had a large influence on the prediction.

4.4 Model Fit Assessment

Cross-validation was performed for the variogram modeling and the kriging predictions to assess the fit of the models. Each site had an associated hold-one-out error, which was the prediction made for that site without that site included in the model minus the true value for that site. Large hold-one-out errors signified that a site's value had a large impact on the predicted surface. In these data, many outlying sites had large hold-one-out errors. This was the result of the lack of nearby sites. This procedure was automated for simple and universal kriging in *GSLIB: Geostatistical Software Library and User's Guide* (Deutsch and Journel 2000), but not for inverse distance weighting where the cross validation was performed by repeated analyses. The cross-validation results were summarized as follows:

- Mean error
- Median of the errors as a percent of the truth
- Mean square error (MSE)
- Square root of the reduced MSE (RedRootMSE) where the reduced MSE was the mean of the squared error divided by the kriging variance.

In a model that fits the data well, the mean error was small as was the error as percent of truth and the MSE. The square root of the reduced MSE was near unity in a model that fit well, and within the limits $1 \pm 2\sqrt{2/n}$ (Magnuson and Sondrup 1998).

5. RESULTS

The results for each data set are presented separately. Data transformation results are presented first, followed by results of the variogram modeling. A comparison of predictions from the three models is presented next. The predictions and standard errors are presented as maps in Appendix B, one for each of the four grids.

5.1 Lithology Data Set

Full quadratic regressions on easting and northing were run for each variable to determine the need for estimating an unknown mean function (see Table 2). Almost all variables were related to at least one of the terms, and many were related to the interaction or quadratic terms as well. Only significant terms were used in exploring isotropy and in universal kriging.

All variables were found to be isotropic after considering directional variograms at four directions. The raw data, as well as residuals of the regressions on latitude and longitude, were considered for isotropy. Figure 5 shows all directional variograms for the C basalt layer elevation for the raw data and the residuals from regression. This example represents the situation for other variables in that there was little difference in the shape, nugget, sill, or range for any direction or for the residuals versus the raw data.

About half of the 39 variables required to produce the 18 data setups were fit with a Gaussian model and half with a spherical model. The spatial ranges of the variables were between 107 to 914 m (350 and 3,000 ft) with most of the small ranges being for interbed soil thickness layers. The variograms were restricted to pairs of sites within 914 m (3,000 ft) because pairs of sites beyond that range were rare. The final data set (discussed next) had variograms (see Table 3 and Appendix B, Figure B-1) that were somewhat consistent. The elevation variables had fairly large ranges (396 to 883 m [1,300 to 2,900 ft]) and the thickness variables had smaller ranges (183 to 259 m [600 to 850 ft]) except the B-C interbed soil thickness, which had a range of 914 m (3,000 ft).

The data, as recorded with gaps, truncated values, and a mix of thickness and elevation, produced acceptable results with the least amount of data manipulation (see Table 4). The three methods for calculating elevation were similar; therefore, it did not matter whether variables were transformed to elevation, thickness, or left as a combination. However, the three methods used to deal with gaps between layers yielded different results. Although the methods that differentiated soil merges and basalt merges were more appealing theoretically, the differences among method results were quite small. The differences were consistent between Methods 2 and 3 (which dealt with gaps and pinches differently than the default method [Method 1]), and Method 3 (which used half the thickness for each layer) predicted higher elevations, as expected. The differences between Methods 1 and 3 were inconsistent across layers, and the only sizable difference between Methods 1 and 2 was the A-B interbed soil elevation. These large differences occurred, as expected, at areas with gaps in the basalt layer (see Figure 6). Though the differences were large (up to 4.6 m [15 ft] in some areas), the predictions were not what the hydrologists expected to see. Because the hydrologists were accustomed to data and maps produced from the default method, and they were the interpreters, the default method was used in the final data set. Some differences also existed between the methods to deal with truncated values. In the C-D interbed soil layer thickness (where most of the truncated data occurred), the estimated values produced a smoother predictive surface. Because the hydrologists preferred the rougher surfaces, truncated values were used. In the A-B interbed soil layer thickness, the one estimated value actually produced a very large cross-validation prediction error. The prediction at this point was a lot larger than the truncated value and any other value for that layer, which was another reason to use the raw truncated data.

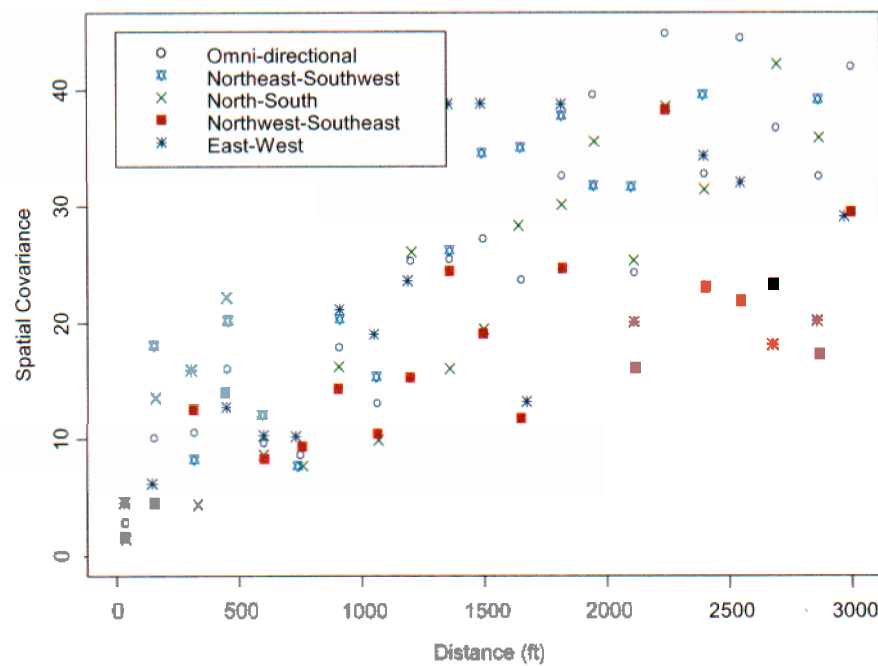
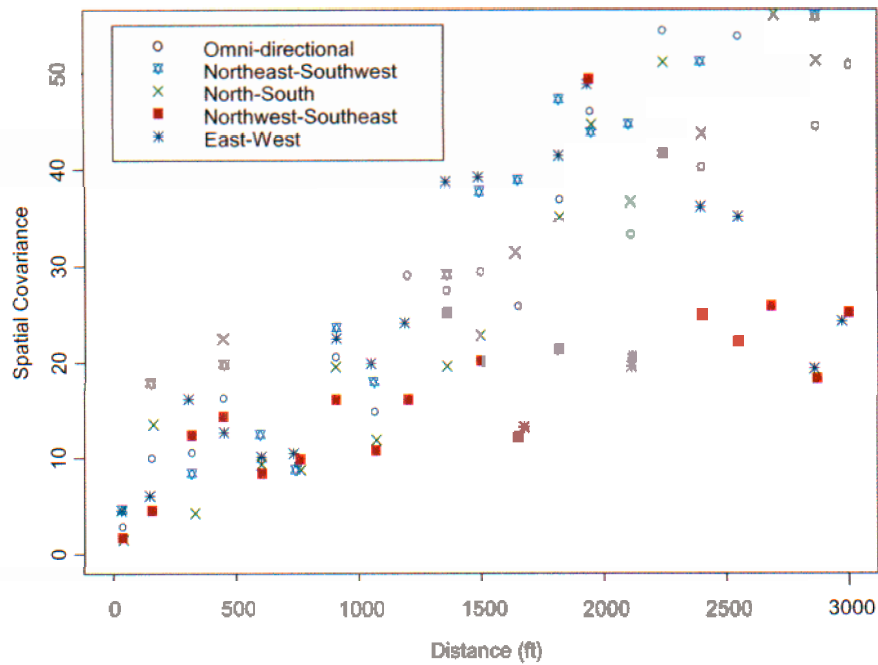


Figure 5. Directional variograms for the C basalt layer raw data (top) and residuals from regression on easting and easting² (bottom). (Gamma is a special covariance and measurements are reported in feet.)

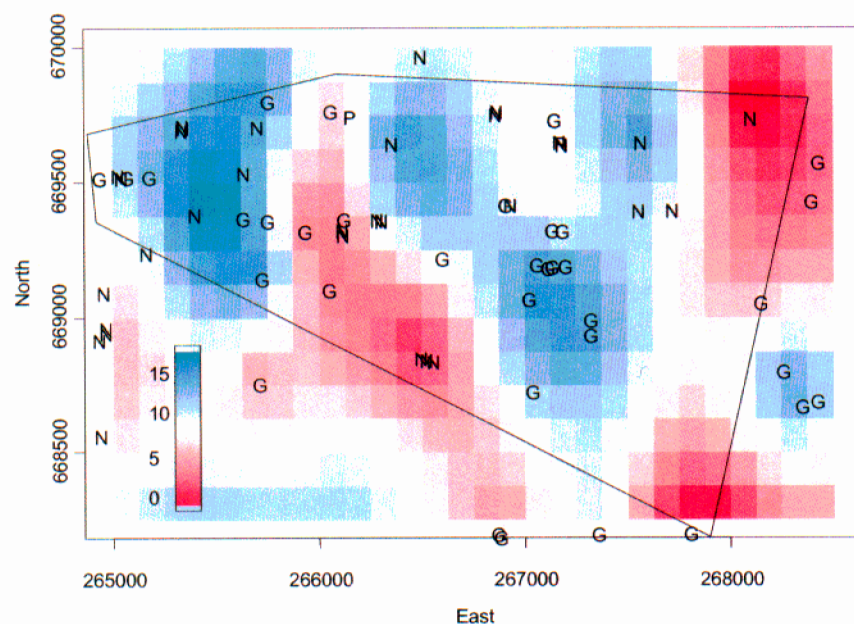


Figure 6. Difference between prediction of bottom elevation of A-B interbed soil from Methods 1 and 2 for dealing with gaps. The G indicates a site location with an A basalt layer gap, P indicates a pinch off of the A-B interbed soil layer, and N indicates neither situation occurred there.

Table 3. Interpolation method and variogram parameters (where appropriate) used to model basalt layer elevations, interbed soil thickness, soil permeability and porosity, and aquifer permeability.

Variable	Nugget	Sill	Range (ft)	Interpolation Method
Surface elevation	0 ft ²	25 ft ²	1,300	Universal kriging
Surficial soil thickness	5 ft ²	44 ft ²	600	Simple kriging
A basalt elevation	0 ft ²	170 ft ²	1,800	Simple kriging
A-B interbed soil thickness	1 ft ²	4 ft ²	850	Universal kriging
B basalt elevation	5 ft ²	31 ft ²	2,900	Universal kriging
B-C interbed soil thickness	0 ft ²	120 ft ²	3,000	Simple kriging
C basalt elevation	0 ft ²	45 ft ²	2,200	Universal kriging
C-D interbed soil thickness	40 ft ²	50 ft ²	700	Universal kriging
B-C interbed porosity	0%	60% ²	1,000	Universal kriging
C-D interbed porosity	0%	30% ²	1,000	Universal kriging
B-C interbed permeability	—	—	—	Inverse distance weighting
C-D interbed porosity	—	—	—	Inverse distance weighting
Transmissivity	0 (ft ² /day) ²	15 (ft ² /day) ²	1,800	Universal kriging

Table 4. Comparison of data evaluation techniques for characterizing lithologic layers (units in feet).

Mean Difference Between Predictions for Different Data Methods (Mean Absolute Difference) ^{a, b}							
Layer	Truncated Values	Data Units			Gaps		
	Prediction of Truncated Value Minus Prediction of Estimated Value	Prediction of Mixed Elevation and Thickness Minus Prediction of Elevation	Prediction of Mixed Elevation and Thickness Minus Prediction of Thickness	Prediction of Elevation Minus Prediction of Thickness	Prediction of Method 1 Minus Prediction of Method 2	Prediction of Method 1 Minus Prediction of Method 3	Prediction of Method 2 Minus Prediction of Method 3
Surface elevation	NA	NA	NA	NA	NA	NA	NA
Surficial soil elevation	NA	-0.11 (1.37)	0	0.11 (1.37)	0	-3.15 (3.15)	-3.15 (3.15)
A basalt elevation	NA	-0.13 (0.18)	-1.02 (1.02)	-1.15 (1.15)	NA	NA	NA
A-B interbed soil elevation	0.27 (0.31)	0.28 (1.09)	-0.88 (0.88)	-1.16 (1.5)	6.71 (6.71)	3.2 (3.2)	-3.52 (3.52)
B basalt elevation	NA	0	-1.11 (1.2)	-1.11 (1.2)	NA	NA	NA
B-C interbed soil elevation	1.86 (1.86)	0.34 (0.57)	-1.11 (1.2)	-1.46 (1.46)	No gaps		
C basalt elevation	NA	0	0.08 (1.3)	0.08 (1.3)	NA	NA	NA
C-D interbed soil elevation	8.17 (8.17)	0.48 (0.87)	0.08 (1.3)	-0.41 (0.98)	0.71 (0.71)	0	-0.71 (0.71)

a. Mean differences, as well as mean absolute differences, are reported to indicate whether the differences were always one-way.

b. NA indicates that the data method did not affect that layer.

The modeling techniques produced very different maps with the kriging and universal kriging results being preferred over inverse distance weighting in general because reliable estimates of variance are available. The inverse distance-weighting surface was sometimes similar to the kriging estimates, but most of the time had small patches of small or large values surrounding observations. The results of the cross-validation exercise revealed that some layers were more difficult to predict in general, and some sites were difficult to predict in some or all layers. The interbed soil thickness layers had large errors as a percent of the truth (see Table 5). This resulted in part from the smaller range of values for thickness. An error of a few feet was larger relative to a thickness of 3 to 6 m (10 to 20 ft) than for an elevation of 1,524 m (5,000 ft). Four layers had extreme RedRootMSE values: A basalt, A-B interbed soil, B-C interbed soil, and C basalt. These were not caused by exceptionally large errors, but because of small kriging variances (see Appendix B, Figure B-2). The sites that were difficult to predict in many of the layers were NA89-2, O-7, and M14S. Two of the sites, NA89-2 and M14S, were isolated sites near the border of the prediction grid; therefore, these large deviations were expected. Site O-7 was interior but far away from its neighbors (compared to other interior sites). Therefore, larger deviations were expected in this area as well.

The kriging standard deviations have similar patterns for many layers as a result of similar variogram range and sill (see Appendix B, Figure B-2). Layers with smaller variogram range generally have a smaller difference between the nugget (i.e., variability among close sites) and sill (i.e., variability among independent sites); therefore, these layers have more constant-looking kriging standard deviation maps (e.g., subsurface, A-B, and C-D interbed soil thickness). Layers with a larger variogram range generally have a larger range of kriging variance caused by the large difference between nugget and sill, and this results in standard deviation maps with more contours (e.g., surface and basalt elevations and B-C interbed soil thickness). Kriging standard deviations are smaller in areas with many sample sites and larger away from the central area where sampling becomes sparser. Grid Refinement 3 also displays small standard deviation areas around clumps of sample sites.

Table 5. Cross-validation statistics for lithology data.^a

Layer	Mean Error		Error as Percent of Truth (Median)		Mean Square Error		RedRootMSE	
	Grid 0	Grid 3	Grid 0	Grid 3	Grid 0	Grid 3	Grid 0	Grid 3
Surface elevation	-0.14	0.04	0	0	18	4	1.2	1.1
Surficial soil thickness	0.25	-0.28	5	-10	21	23	0.9	1.0
A basalt elevation	-0.69	-0.76	0	0	50	46	2.0	2.2
A-B interbed soil thickness	-0.01	0.24	-29	-2	10	5	2.1	1.5
B basalt elevation	-0.06	-0.34	0	0	33	12	1.4	1.2
B-C interbed soil thickness	-0.35	-0.27	-5	-5	78	81	4.8	6.3
C basalt elevation	-0.35	-0.46	0	0	79	12	2.3	1.7
C-D interbed soil thickness	-0.27	-0.10	-10	-10	4	112	1.0	1.2

a. Cross validation was done for predictions made for the base grid (only for sites within the base grid) and Grid 3 (only for sites inside Grid 3). Values of RedRootMSE in **bold** font exceed the limit given in the text.

5.2 Hydraulic Characteristics Data Set

The porosity data were related to easting, northing, and easting x northing in the B-C interbed layer and only easting in the C-D interbed layer (see Table 2). The residuals of these regressions showed no signs of anisotropy. The variograms were modeled at a maximum distance of 762 and 914 m (2,500 and 3,000 ft) for B-C and C-D interbed layers, respectively. The classical estimator was used and a lag of 91 and 61 m (300 and 200 ft) was specified for B-C and C-D, respectively. Site I4-S was influential in variogram modeling for the B-C interbed. It was located within the SDA, but was on the outer edge of the convex data hull. A measurement was taken at only one depth at this site, which had the largest mean porosity of any site in the B-C interbed. Those sites near I4-S had porosity values approximately 30% less than the reported 47.5 %. Without this site, the variogram was fit with a spherical model with no nugget, a sill of 5.6 m^2 (60 ft^2), and a range of 305 m (1,000 ft) (see Table 3 and Appendix B, Figure B-1[i]). The C-D layer contained no suspect values and was modeled by a spherical model with no nugget, a sill of 2.8 m^2 (30 ft^2), and a range of 305 m (1,000 ft) (see Table 3 and Appendix B, Figure B-1[k]).

The variograms for porosity showed the smallest distances at about 152 m (500 ft) and a definite increase came from that area, but the data did not extend to the origin for modeling. Therefore, an assumption was made about the small-scale spatial variability in the data. It was assumed that the data should be modeled with a spherical variogram, which is more conservative than the Gaussian model, because the spatial correlation was assumed to decrease more rapidly as the distance increased.

The cross-validation statistics for porosity indicated a good fit for both layers (see Table 6). The mean error, error as percent of truth, and MSE are similar for the base grid and Grid 3, and also between the layers. The reduced root MSE values, though similar for the two grids, were dissimilar between the layers. The B-C layer had a reduced root MSE that was above the limit and more than twice for the value from the C-D layer. This resulted from the lower number of sites in the B-C layer (17 versus 24 in the C-D layer) and from the higher variability in the B-C layer (B-C sill = 5.6 m^2 [60 ft^2] and C-D sill = 2.8 m^2 [30 ft^2]) (see Appendix B, Figure B-1 [i, k] and Table 3). The kriging standard deviation maps (see Appendix B, Figure B-2) display an increasing magnitude at an increasing distance from the cluster of sample sites.

The permeability data followed a lognormal distribution; therefore, a natural log transformation was necessary for kriging. Kriging methods were somewhat robust to nonnormality, but not to nonsymmetry. The permeability data were related to neither northing nor easting for the B-C interbed, but were related to northing and northing² for the C-D interbed (see Table 2). The variograms were difficult to model because of a lack of site locations at small distances (see Appendix B, Figure B-1 [l, m]). The small-scale spatial variation was assumed instead of modeled. To interpret the kriging predictions, the values were back-transformed, including the bias correction. The kriging variance was quite large because of the small number of close stations. Therefore, the back-transformed data were large and out of the range of the data values themselves. Kriging was an inappropriate method for these data. Two alternative models for permeability were investigated. A normal-score transformation resulted in similarly unreasonable predictions because the variogram was pure nugget with no visible range. The inverse distance-weighting method predictions were reasonable (see Appendix B, Figure B-3) relative to the range of values observed.

The cross-validation statistics provided conflicting results (see Table 6). Those measures that were not variance adjusted (i.e., mean error, error as percent of truth, and MSE) were extremely large. Though the variance adjusted RedRootMSE was larger than the given limit for the C-D layer, it was not as extreme as the other measures.

Because the standard deviation for inverse distance weighting uses the same weights as the predictions, the patterns are similar (see Appendix B, Figure B-2). The standard deviation is smaller in

Table 6. Cross validation statistics for hydraulic characteristics.^a

Variable	Layer	Mean Error		Error as Percent of Truth		Mean Square Error		RedRootMSE	
		Grid 0	Grid 3	Grid 0	Grid 3	Grid 0	Grid 3	Grid 0	Grid 3
Porosity	B-C	-0.43	-0.43	3	3	122	122	2.2	2.2
	C-D	0.47	-0.13	1	1	40	40	1	1
Permeability	B-C	-41	NA	1,079	NA	2,507,802	NA	1.6	NA
	C-D	62	NA	841	NA	2,685,023	NA	2.6	NA

a. Porosity was modeled using universal kriging and permeability was modeled using inverse distance weighting. The inverse distance weighting was done without respect to grid; therefore, general statistics are listed under Grid 0. Values of RedRootMSE in **bold** font exceed the limit given in the text.

NA = not applicable

areas of dense sampling, but these clusters are less defined for the inverse distance weighting than for kriging. This results from the distance variance relationship, which in general is steeper for inverse distance weighting, causing the variance to be influenced by close sites and giving all sites past a specific distance approximately equal weight. The variogram relationship generally has a larger range of influence and the weights drop off more slowly as distance increases.

5.3 Aquifer Transmissivity Data Set

Aquifer transmissivity, log transformed, was significantly related to easting and northing (see Table 2). The directional variograms of the raw data and the residuals were similar, indicating isotropic data.

With limited short-range data, the variogram cloud was used to verify the chosen variogram model (see Table 3 and Appendix B, Figure B-1[m]). Because the small-scale results were similar for the cloud and the model, moderate confidence was placed in the model. The universal kriging predictions were compared to inverse distance weighting to determine the utility of the variogram model. The inverse distance weighting results were not realistic. The majority of the area was predicted to be very high with small pockets of small values around observations.

The back-transformation for aquifer transmissivity, which was lognormal, was successfully applied here compared to the unsuccessful attempt for interbed soil permeability. The spatial continuity was greater than the interbed soil permeability. The range was estimated at 549 m (1,800 ft) for the aquifer transmissivity but 305 m (1,000 ft) for the soil permeability. Therefore, even though relatively few samples were taken at close distances, those that were taken fully support the assumed model. The stable variogram also explains the moderate-sized kriging variance, hence the utility of the back-transformation for these data.

The cross-validation statistics also indicated a good fitting model (see Table 7). The unadjusted error measurements were small and the reduced root MSE was within the limits for the sample size. The kriging standard deviation pattern followed that of the kriging predictions because of the small number of observations (n = 21) (see Appendix B, Figure B-2).

Table 7. Cross validation statistics for aquifer permeability based on log-transformed data.

Mean Error		Error as Percent of Truth		Mean Square Error		RedRootMSE	
Grid 0	Grid 3	Grid 0	Grid 3	Grid 0	Grid 3	Grid 0	Grid 3
0.11	0.11	1	1	4	4	0.75	0.75

6. DISCUSSION

The models provide predictions over the area of interest at the density needed for inclusion in a hydrologic model and do so with reasonable precision. The prediction maps (see Appendix B, Figure B-3) reveal some interesting relationships. Many of the lithology layers show similarities in shape of contours, though not actual values. The bottom elevation of the interbed soil layer has roughly the same shape as the bottom elevation of the basalt layer above it (or surface layer for the surficial soils). This relationship breaks down somewhat at increasing depth and for increasing grid density. For example, the C-D interbed soil layer only looks similar to the C basalt layer for the base grid, but not for the other grids. The B-C interbed soil layer looks similar to the B basalt layer for all but the most intense grid (Grid 3). A general similarity of shape also is shown for the surface elevation down to the B basalt layer for all grids. The prediction contours for soil permeability also are similar between the B-C and C-D interbed soil layers for all grids. The prediction contours for porosity of those soil layers show increasing differences as the grid intensity increases.

6.1 Precision of the Predictions and Fit of the Model

The kriging standard deviations indicate that the precision of the predictions is good overall, given the small number of sample site locations (see Table 8 and Appendix B, Figure B-2). The standard errors need to be taken into account when using the predictions. Predicting with hydrologic models using these predictions should take into account the variability in these predictions as well as the variability in the future model. For predictions that have a large coefficient of variation (i.e., $CV = \text{standard deviation} / \text{mean}$), the end predictions may not be deemed useful because of their combined model variance. The modeling standard deviations are given in Table 8, in addition to the CVs. A CV greater than 10% is often considered large. None of the lithology predictions has a large CV. The CVs based on elevation appear to be small, but the predictions used for mapping and in other models are in elevation; therefore, they are appropriate. The other variables have larger CVs. This results from the small number of site locations in these data sets. The location of future modeling also is a factor. The prediction standard deviation is smaller in areas of dense sampling; therefore, future modeling also will be more precise in these areas.

Table 8. Modeling standard error for Grid Refinement 2.^a

Data Set	Variable	Minimum Standard Error	Maximum Standard Error	Mean Standard Error	Mean Coefficient of Variation (%)
Lithology	Surface elevation	0.7	5.1	3.2	0.06
	Surficial soil thickness	3.0	7.1	6.2	0.12
	A basalt elevation	1.7	12.6	7.0	0.14
	A-B interbed soil thickness	1.2	2.3	1.9	0.04
	B basalt elevation	2.5	5.1	3.5	0.07
	B-C interbed soil thickness	1.1	8.5	4.6	0.09
	C basalt elevation	0.8	6.0	3.4	0.07
	C-D interbed soil thickness	7.2	9.6	9.0	0.19
Interbed hydraulic	B-C interbed porosity	2.5	9.0	7.1	22.0

Table 8. (continued).

Data Set	Variable	Minimum Standard Error	Maximum Standard Error	Mean Standard Error	Mean Coefficient of Variation (%)
	C-D interbed porosity	1.6	6.1	4.8	11.0
	B-C interbed permeability	4.3	1,342	296	25.0
	C-D interbed permeability	4.6	1,023	138	21.0
Aquifer permeability	Transmissivity	0.3	889,982	24,885	48.0

a. The coefficient of variation (i.e., CV = standard error [standard deviation of the prediction] divided by the prediction) for lithology data is based on values converted to elevations because the small thickness values bias the results.

6.2 Cross Validation

Another indication of the fit of these models is the cross validation, which gives different results for different variables. Five of the eight lithology layers have large cross-validation statistics (see Table 5). These five lithology layers are listed below:

- A basalt
- A-B interbed soil
- B basalt
- B-C interbed soil
- C basalt.

These large cross-validation statistics are generally caused by a few sites that have large influences because of their locations. Sites that are far away from other sites are very influential in the predictive surface because, if their value is different from the closest sites (even if they are far away), then the cross-validation error will be large. The large cross-validation errors for soil porosity in the B-C interbed soil and permeability in the C-D interbed result from the small sample size, as well as the site locations (see Table 6).

6.3 Sparseness of Data

Problems in modeling these data include sparseness as well as large distances between many of the sites. These problems affected all modeling techniques; however, it mostly affected kriging where sparse data could not be used to reliably estimate the semivariance at small distances. The variogram model is most dependent on the fitting at small distances because that is where the differences occur relative to model shape and nugget. Beyond the range, the variogram is of little interest. Therefore, the pairs of points within the small distances are important. The variogram model parameters are used to predict values at unsampled locations as well as to estimate the variance of the predictions; therefore, these also are most affected by the variogram model at small ranges. Because the weights in the kriging equations change (sometimes drastically) from close distances to the range, these values can greatly affect the predictions. The predictions also will vary depending on the form of the model at small distances. For instance, the Gaussian and spherical models have different shapes at small distances. The Gaussian model starts horizontally and then slopes upward, producing a surface that is quite smooth because close sites

are predicted to be similar. The spherical model starts with an upward slope, producing a surface that is rougher because close sites can be predicted to show more difference. Because of these reasons, it is essential that sites in the area of interest be close together. Sparse sites will produce predictions that may be biased because of inaccurate variogram modeling as well as predictions that have large variances; therefore, little confidence is placed in those values.

6.4 Improvements to Predictions

Improvements can be made to these predictions without collecting further information by (1) using simulation rather than model prediction and by (2) using covariates to improve prediction. Simulation methods result in a distribution of predictions for each grid location instead of a single number. Simulation results provide a broader picture because many of the predictions are realized. Simulation also may provide a way to simultaneously predict values for the various lithology layers, which could improve overall prediction. Implementing simulation methods is more time and computer intensive than kriging models and has not been explored with these data.

The second improvement to predictions may come in the form of using related variables. Cokriging (i.e., kriging plus a covariate) is one approach to incorporate more information. In order for the covariate to be useful, it should be related to the variable of interest and should be measured at more locations than the variable of interest. Covariates that do not meet these criteria could unduly increase the kriging variance without changing the predictions significantly. Simply using upper layers in the lithology data for cokriging might not be very useful because there are no extra sites. Another weakness in using the lithology data as covariates is that cokriging makes an additional assumption of a linear model of coregionalization. This assumption is generally satisfied by having variables and covariates that have similar structure (i.e., model shape) and spatial range. This is not the case for the lithology data. However, new approaches could be explored for dealing with this. Another approach would be to predict layers one at a time and use the predictions for the next lower layer. This could start by using a virtually continuous surface elevation map to help predict surficial soil depth.

7. REFERENCES

- 42 USC § 9601 et seq., 1980, *United States Code*, “Comprehensive Environmental Response, Compensation and Liability Act of 1980 (CERCLA/Superfund).”
- Anderson, S. R., D. A. Ackerman, M. J. Liszewski, and R. M. Freiburger, 1996, *Stratigraphic Data for Wells At and Near the Idaho National Engineering Laboratory, Idaho*, DOE/ID-22127, Open-File Report 96-248, U.S. Geological Survey, Idaho Falls, Idaho.
- Barracough, J. T., J. B. Robertson, V. J. Janzer, and L. G. Saindon, 1976, *Hydrology of the Solid Waste Burial Ground, as Related to the Potential Migration of Radionuclides, Idaho National Engineering Laboratory*, IDO-22056, Open File Report 76-471, U.S. Geological Survey, Idaho Falls, Idaho.
- Cressie, N. A. C., 1993, *Statistics for Spatial Data*, New York: John Wiley & Sons.
- Deutsch, C. V., and A. G. Journel, 2000, *GSLIB: Geostatistical Software Library and User's Guide*, New York: Oxford University Press.
- DOE-ID, 1991, *Federal Facility Agreement and Consent Order for the Idaho National Engineering Laboratory*, U.S. Department of Energy Idaho Operations Office; U.S. Environmental Protection Agency, Region 10; Idaho Department of Health and Welfare.
- Freeze, R. A., and J. A. Cherry, 1979, *Groundwater*, New Jersey: Prentice-Hall.
- Galouchko, V., 2002, *3D Field*, URL: <http://field.hypermart.net/>, Web page visited March 18, 2002.
- Magnuson, S. O., and A. J. Sondrup, 1998, *Development, Calibration, and Predictive Results of a Simulator for Subsurface Pathway Fate and Transport of Aqueous- and Gaseous-Phase Contaminants in the Subsurface Disposal Area at the Idaho National Engineering and Environmental Laboratory*, INEEL/EXT-097-00609, Idaho National Engineering and Environmental Laboratory, Lockheed Martin Idaho Technologies Company, Idaho Falls, Idaho.
- MathSoft, 2000, *S+SPATIALSTATS User's Manual Version 1.5*, Data Analysis Products Division, MathSoft, Seattle, Washington.
- McElroy, D. L., and J. M. Hubbell, 1990, *Hydrologic and Physical Properties of Sediments at the Radioactive Waste Management Complex*, EGG-BG-9147, Idaho National Engineering and Environmental Laboratory, EG&G Idaho, Idaho Falls, Idaho.
- Reich, Robin M., and Richard Davis, 2000, *Manual – Quantitative Spatial Analysis (Course Notes NR/ST 523)*, Colorado State University, Fort Collins, Colorado, URL: <http://www.cnr.colosate.edu/~robin/spatmast.pdf>, Web page visited March 18, 2002.
- USGS, 2000, *Measurement of Hydraulic Properties of the B-C Interbed and Their Influence on Contaminant Transport in the Unsaturated Zone at the Idaho National Engineering and Environmental Laboratory, Idaho*, Water-Resources Investigation Report 00-4073, U.S. Geological Survey.

- Wylie, A. H., 1996, *Pumping Test of Pit 9 Production Wells*, Engineering Design File INEL-96/171, Idaho National Engineering and Environmental Laboratory, Lockheed Martin Idaho Technologies Company, Idaho Falls, Idaho.
- Wylie, A. H., and J. M. Hubbell, 1994, *Aquifer Testing of Wells M1S, M3S, M4D, M6S, M7S, and M10S at the Radioactive Waste Management Complex*, Engineering Design File ER-WAG7-26, Rev. 1, Idaho National Engineering and Environmental Laboratory, EG&G Idaho, Idaho Falls, Idaho.